University of West Bohemia

Faculty of Applied Sciences

Department of Mathematics

BACHELOR'S THESIS

Union-closed sets conjecture and the strength of constraints in linear programming

2024

Jakub Koňařík

Abstract

In the present thesis we familiarize the reader with the Union-closed sets conjecture. The conjecture states that any finite union-closed family of sets has an element that belongs to at least half of the member sets. We examine in detail the conjecture's assumptions, known observations, equivalent formulations, selected known partial results and results regarding small families of sets. We also mention recent breakthroughs obtaining a constant fraction for the conjecture.

The main result presented in this thesis is that the Union-closed sets conjecture holds when the size of the biggest set is at most 14, we improve upon the previously known results for sizes 7,9,10,11,12 by Poonen [20], Lo Faro [11], Marković [18], Bošnjak and Marković [3] and Vučković and Živković [28] respectively. This main result was obtained in a joint project of the author of the thesis and Jana Chrastinová, Adam Kabela and Jakub Teska. This thesis provides a detailed introduction and explanation of the methods used as well as a summary of our computer assisted proof.

Our approach is to formulate a restriction of the Union-closed sets conjecture as an integer linear programming problem. For comupting purposes, we relax the problem to a linear programming problem. We compensate for the relaxation by splitting the proof into multiple cases, iteratively proving ever stronger inequalities and adding them as constraints. We use hypergraph isomorphism for organising the case hierarchy. The proof is computer-assisted.

Key words

Union-closed sets conjecture, Small families of sets, Linear programming, LP Relaxation, LP Duality, Computer assisted proof.

Abstrakt

V předkládané práci je čtenář nejprve seznámen s Franklovou hypotézou. Franklova hypotéza říká, že všechny konečné systémy množin uzavřené na sjednocení obsahují prvek, který patří do alespoň poloviny množin v systému. V práci detailně rozebíráme předpoklady hypotézy, základní poznatky, ekvivalentní formulace, vybrané známé částečné výsledky a výsledky týkající se malých systémů množin. Zmíněn je též nedávný průlom v nalezení konstantního dolního odhadu hypotézy.

Hlavní výsledek prezentovaný v této práci je důkaz, že Franklova hypotéza platí pro systémy množin uzavřených na sjednocení, kde velikost největší množiny v systému je nejvýše 14. Tento výsledek zlepšuje doposud známé výsledky pro velikosti 7,9,10,11,12 od Poonen [20], Lo Faro [11], Marković [18], Bošnjak and Marković [3] a Vučković a Živković [28] v tomto pořadí. Důkaz je výsledkem skupinového projektu autora této práce, Jany Chrastinové, Adama Kabely a Jakuba Tesky. Předkládaná práce obsahuje detailní představení a vysvětlení použitých metod a také shrnutí celého počítačem asistovaného důkazu.

V důkazu nejdříve formulujeme restrikci Franklovy hypotézy jako problém celočíselného lineárního programování, který z výpočetních důvodů relaxujeme na problém lineárního programování. Relaxaci kompenzujeme rozdělením důkazu do několika částí, iterativně dokazujeme čím dál silnější nerovnice, které přidáváme do programu jako podmínky. K organizaci a řazení částí důkazu využíváme izomorfizmus hypergrafů.

Klíčová slova

Franklova hypotéza o množinových systémech uzavřených na sjednocení, malé systémy množin, lineární programování, LP relaxace, LP dualita, důkaz pomocí počítače.

Acknowledgement

I would like to thank Adam Kabela for supervising my thesis. I would also like to thank Jana Chrastinová, Adam Kabela and Jakub Teska for the joint work on a research project focused on the Union-closed sets conjecture.

Poděkování

Touto cestou bych rád poděkoval Adamu Kabelovi za vedení mé práce. Také bych rád poděkoval Janě Chrastinové, Adamu Kabelovi a Jakubu Teskovi za spolupráci při výzkumu na téma Franklovy hypotézy.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a za použití literatury, jejíž úplný seznam je uveden na konci práce.

V Plzni, dne ....

....

Jakub Koňařík

Contents

Notation

$\mathcal{F}, \mathcal{G}, \mathcal{H}, \ldots$	Families of sets
$A, B, C, \ldots$	Sets, member sets
$a, b, c, \ldots$	Elements of member sets
$\mathcal{F}^c$	The family of complements of the member sets of $\mathcal{F}$
2^A	Power set of a set A
k-set	A set of size k
$\boldsymbol{A}, \boldsymbol{B}, \ldots$	Matrices
$\boldsymbol{a}, \boldsymbol{b}, \ldots$	Vectors
$\boldsymbol{A}^T$	Transpose of a matrix A
$\boldsymbol{c}^T$	Transpose of a vector c
$G, \ldots$	Graphs
$(L, \preccurlyeq), \ldots$	Lattices
$\mathbb{R}$, $\mathbb{Z}$	Real and integer domains
$\mathbb{R}^n$, $\mathbb{Z}^n$	Domains of real and integer n-dimensional vectors
$\mathbb{R}^{m \times n}$, $\mathbb{Z}^{m \times n}$	Domains of real and integer m by n matrices

Elementary definitions

Family of sets	A *family of sets* is a collection of pairwise distinct sets.
Union-closed	A family of sets $\mathcal{F}$ is *union-closed* if for every $A, B \in \mathcal{F}$ the union $A \cup B$ belongs to $\mathcal{F}$.
Universe	Let $\mathcal{F}$ be a family of sets. The union of all sets in $\mathcal{F}$ is called the universe of $\mathcal{F}$, denoted by $U(\mathcal{F})$.
Abundant, rare element, frequency	Let $\mathcal{F}$ be a family of sets. An element $e \in U(\mathcal{F})$ is called abundant if it belongs to at least half of the member sets of $\mathcal{F}$ and *rare* if it belongs to at most half of the member sets of $\mathcal{F}$. The frequency of e is the number of set in $\mathcal{F}$ which contain e.
Bijection	Let A, B be sets and let $f : A \to B$. We say that f is a *bijection* if $\forall a, b \in A : f(a) = f(b) \Rightarrow a = b$ and $\forall b \in B \; \exists a \in A : f(a) = b$ and also $\forall a \in A \exists b \in B : f(a) = b$.
Nonoriented graph	Let V be a finite set and let E be a set of 2-subsets of V. Then the tuple $(V, E) = G$ is called a *nonoriented graph*, or *graph* for short. Vertices $v_1, v_2 \in V$ are adjacent in G if and only if $\{v_1, v_2\} \in E$
Stable set	Let $G = (V, E)$ be a graph. A subset of the vertex set $S \subseteq V$ is called stable if no two vertices in S are adjacent in G.
Bipartite graph	Bipartite graph is a graph whose vertex set V can be divided into two disjoint stable sets.
Binary relation	A subset ρ of the cartesian product $X \times X$ is called a binary relation on X. We say that ρ is reflexive if $\forall x \in X : (x, x) \in \rho$, antisymmetric if $\forall x, y \in X : (x, y) \in \rho$ and $(y, x) \in \rho \Rightarrow x = y$ and transitive if $\forall x, y, z \in X : (x, y), (y, z) \in \rho \Rightarrow (x, z) \in \rho$.
Lattice	Let L be a set and $\preccurlyeq$ be a reflexive, transitive and antisymmetric binary relation on L. We define the greatest lower bound $a \wedge b$ to be $x \in L$ such that $x \preccurlyeq a, \; x \preccurlyeq b$ and also if for some$z$ it holds $z \preccurlyeq a, \; z \preccurlyeq a$ then $z \preccurlyeq x$. The least upper bound $a \vee b$ is $y \in L$ such that $x \succcurlyeq a, \; x \succcurlyeq b$ and if for some $z \in L$ it holds $z \succcurlyeq a, \; z \succcurlyeq a$ then $z \succcurlyeq y$. If $\forall a, b \in L \; \exists a \wedge b \; \exists a \vee b$, then $(L, \preccurlyeq)$ is called a lattice.

1 Introduction

The aims of this thesis are to familiarize the reader with the Union-closed sets conjecture and related partial results and also to examine the application of linear programming to a restriction of the conjecture.

The text is divided into two chapters. The first chapter first briefly introduces the Union-closed sets conjecture in Section 1.1. Afterwards, we discuss the assumptions of the conjecture in Section 1.2, examine the conjecture's equivalent formulations in Section 1.4 and summarize known partial results in Sections 1.5, 1.6 and 1.7. We give a brief explanation of linear programming in Section 1.8.

In Chapter 2 we apply linear programming to a restriction of the Union-closed sets conjecture where the size of the biggest set is bounded. We prove that the conjecture holds for families on at most 14 elements. We give a detailed explanation of our proof architecture in Sections 2.3 and 2.4. Afterwards, we provide an outline of our proof and discuss corollaries of our results.

1.1 Union-closed sets conjecture

The Union-closed sets conjecture is a famous conjecture in combinatorics. The conjecture is commonly attributed to Frankl [13] and dated to 1979, but Balla, Bollobás and Eccles [2] mention that the conjecture is even older and was known in the 1970s as a folklore conjecture. A detailed historical and mathematical overview of the conjecture can be found in the survey by Bruhn and Schaudt [4] titled *The journey of the Union-closed sets conjecture* describing the travel of the conjecture from continent to continent in the pre-internet era as well as the journey towards understanding the problem through partial results.

The conjecture concerns families of sets that are closed under taking unions. A *family* of sets is a collection of pairwise distinct sets. A family $\mathcal{F}$ is *union-closed* if for every two member-sets $A, B \in \mathcal{F}$ their union $A \cup B$ also belongs to $\mathcal{F}$. The Union-closed sets conjecture states the following.

Conjecture 1.1.1 (Union-closed sets conjecture): *Any finite union-closed family of sets $\mathcal{F} \neq \{\emptyset\}$ has an element that belongs to at least half of the member-sets of $\mathcal{F}$.*

Despite its simple formulation, the conjecture remains an open problem in mathematics and even proving partial results can be challenging.

1.2 Assumptions of the Union-closed sets conjecture

In this section, we discuss the assumptions of the Union-closed sets conjecture and provide examples showing why every assumption is necessary.

We start with the assumption that families are finite. The conjecture does not hold for infinite union-closed families of sets, as shown, for instance, in [20]. Consider the family consisting of sets $\{n,\ n+1,\ n+2,\ ...\}\ \forall n \in \mathbb{N}$, depicted in Figure 1. This family has infinitely many member sets, but each element appears in finitely many member sets.

$$\begin{array}{r} \{1,\ 2,\ 3,\ 4,\ 5, ...\} \\ \{2,\ 3,\ 4,\ 5, ...\} \\ \{3,\ 4,\ 5, ...\} \\ \{4,\ 5, ...\} \\ \ddots\ \vdots \end{array}$$

Figure 1: A family with infinitely many member sets where every element n appears in precisely n member sets.

The next assumption is union-closure. Figure 2 shows two finite families that are not union-closed and have no abundant elements.

$$\begin{array}{c} \{1,2,3\} \\ \{1\}\ \{2\}\ \{3\} \\ \emptyset \end{array} \qquad \begin{array}{c} \{1,2,3,4\} \\ \{1,2\}\ \{2,3\}\ \{3,4\} \\ \{1\}\ \{2\}\ \{3\}\ \{4\} \\ \emptyset \end{array}$$

Figure 2: Two examples of not-union-closed families with no abundant elements. The frequencies of the most common element are 2/5 and 4/9 respectively.

In the case $\mathcal{F} = \{\emptyset\}$ it is immediate that there are no elements in $U(\mathcal{F})$ and thus there can not be an abundant element.

A family is defined as a collection of pairwise distinct sets. Figure 3 shows a collection that contains all the sets of $2^{\{1,2,3\}}$ and where singleton sets are duplicated. One can easily check that all elements belong to less than half of the member sets.

$$\begin{array}{c} \{1,2,3\} \\ \{1,2\}\ \{2,3\}\ \{1,3\} \\ \{1\}\ \{1\}\ \{2\}\ \{2\}\ \{3\}\ \{3\} \\ \emptyset \end{array}$$

Figure 3: An example of a union-closed collection of sets that is not pairwise-distinct and where each element belongs to 5 of the 11 member sets.

1.3 Elementary observations

First we note that if Conjecture 1.1.1 is true, then the constant of one half is tight. To see this, consider a simple family as follows. For any finite set A, let $\mathcal{F}$ be the family of all subsets of A. It is easy to see that $\mathcal{F}$ is union-closed. Moreover, every element of A appears in exactly half of the member-sets of $\mathcal{F}$ since for every subset $B \in \mathcal{F}$ it's complement $A \setminus B$ also belongs to $\mathcal{F}$ and the complement $A \setminus B$ is unique for every B.

Next simple observation is that a union-closed family always contains it's universe. Union closure guarantees that the union of every two sets A, B in $\mathcal{F}$ also belongs to $\mathcal{F}$. So $A \cup B$ belongs to $\mathcal{F}$ and by union closure $A \cup B \cup C$ also belongs to $\mathcal{F}$ for $C \in \mathcal{F}$. This process continues until we get that $\cup_{A \in \mathcal{F}} A$, the universe of $\mathcal{F}$, also belongs to $\mathcal{F}$.

The following observation talks about families which contain small member sets. It was first observed by Renaud and Sarvate [23].

Observation 1.3.1: *Let $\mathcal{F}$ be a finite union-closed family. If $\mathcal{F}$ contains a singleton member set $\{a\}$, then a is abundant in $\mathcal{F}$. Similarly, if $\mathcal{F}$ contains a member set $\{a, b\}$, then at least one of the elements a, b is abundant in $\mathcal{F}$.*

Proof: Let $\mathcal{F}$ be a union-closed family and let $\{a\} \in \mathcal{F}$. Then consider a member set $A \in \mathcal{F}$, either $a \in A$ or $a \notin A$. In the case that $a \notin A$ by union closure $\mathcal{F}$ contains $\{a\} \cup A$. We note that for any distinct $A, B \in \mathcal{F}$ the sets $A \cup \{a\}$ and $B \cup \{a\}$ are distinct as well. Thus there are at least as many sets containing a as there are sets not containing a and therefore a belongs to at least half of the member sets of $\mathcal{F}$.

Now let $\mathcal{F}$ be a union-closed family and let $\{a, b\} \in \mathcal{F}$. Denote $\mathcal{F}_{ab}$ the family of all member sets of $\mathcal{F}$ that contain both a and b. Similarly, denote $\mathcal{F}_{a\not b}$ and $\mathcal{F}_{\not a b}$ the families of member sets of $\mathcal{F}$ that contain exactly one of a, b and $\mathcal{F}_{\not a \not b}$ the family of member sets of $\mathcal{F}$ containing neither a nor b. We want to show that there must be at least as many sets in $\mathcal{F}_{ab}$ as there are in $\mathcal{F}_{\not a \not b}$. Again, consider a member set $A \in \mathcal{F}_{\not a \not b}$, by union closure $\mathcal{F}$ must contain $\{a, b\} \cup A$ which belongs to $\mathcal{F}_{ab}$. Thus $|\mathcal{F}_{ab}| \geq |\mathcal{F}_{\not a \not b}|$. Now we may, without loss of generality, assume that $|\mathcal{F}_{a\not b}| \geq |\mathcal{F}_{\not a b}|$. Then a belongs to at least half of the member sets in $\mathcal{F}$. $\square$

We note that the pattern of Observation 1.3.1 breaks at sets of size 3. This was first observed by Renaud and Sarvate [23], who constructed a finite union-closed family containing $\{1, 2, 3\}$ such that no elements of $\{1, 2, 3\}$ appear in at least half of the member sets of $\mathcal{F}$. We include a similar construction by Poonen [20] because the original paper [23] seems to be difficult to find online. For the ease of understanding the structure of Poonen's construction $\mathcal{F}$, we consider the family $\mathcal{F}^c$ whose sets are precisely the complements of member sets of $\mathcal{F}$. The family $\mathcal{F}^c$ is depicted in Figure 4. We show that $\mathcal{F}^c$ is intersection-closed and that each element of $\{1, 2, 3\}$ belongs to more than half of the member sets of $\mathcal{F}^c$, which is equivalent to the elements not being abundant in $\mathcal{F}$, as discussed in Section 1.4.

$$\{1,2,3,4,5,6,7,8,9\}$$
$$\{4,5,6,7,8,9\}$$
$$\begin{array}{ccc}
 & \{1,2,3,4\} & \{2,3,4\} \\
\{1,2,3\} & \{1,2,3,5\} & \{2,3,5\} \\
\{1,2\}\ \{2,3\}\ \{1,3\} & \{1,2,3,6\} & \{1,3,6\} \\
\{1\}\ \{2\}\ \{3\} & \{1,2,3,7\} & \{1,3,7\} \\
 & \{1,2,3,8\} & \{1,2,8\} \\
\emptyset & \{1,2,3,9\} & \{1,2,9\}
\end{array}$$
$$\{4\}\ \{5\}\ \{6\}\ \{7\}\ \{8\}\ \{9\}$$

Figure 4: A family $\mathcal{F}^c$ containing $\{1, 2, 3\}$ such that none of the elements of $\{1, 2, 3\}$ appear in at most half of the member sets of $\mathcal{F}^c$

Observation 1.3.2: *The family $\mathcal{F}^c$ depicted in Figure 4 is intersection-closed and each element of $\{1, 2, 3\}$ belongs to more than half of the member sets of $\mathcal{F}^c$.*

Proof: We want to show that $\mathcal{F}^c$ is intersection-closed and that none of the elements of $\{1,2,3\}$ are rare. Clearly $\mathcal{F}^c$ contains $2^{\{1,2,3\}}$ as a sub-family. It is immediate that the intersection of any two 4-sets or of any two 3-sets always belongs to $2^{\{1,2,3\}}$. The same also applies the intersection of a 3-set and a 4-set from $\mathcal{F}^c$, as the intersection is either a 2-subset of $2^{\{1,2,3\}}$, $\{1,2,3\}$ itself or one of the other 3-sets in $\mathcal{F}^c$. Intersecting the only 6-set in $\mathcal{F}^c$ with a subset of $2^{\{1,2,3\}}$ always yields $\emptyset$ and the intersection of the 6-set with any 4-set or 3-set in $\mathcal{F}^c$ is a singleton set all of which belong to $\mathcal{F}^c$. The intersection of the universe with any other set in $\mathcal{F}^c$ is the set itself and the intersection of a singleton set with any other set is either the singleton set itself or the emptyset, all of which are present in $\mathcal{F}^c$. Therefore we conclude that $\mathcal{F}$ is intersection-closed. The reader can easily verify that each element of $\{1,2,3\}$ belongs to 15 out of the 28 member sets and thus none of them is rare. $\square$

Corollary 1.3.2.1: *The existence of a 3-set A in a finite union-closed family $\mathcal{F}$ does not imply that any of the elements of A are abundant in $\mathcal{F}$.*

Another trivial observation is that the Union-closed sets conjecture holds for union-closed families $\mathcal{F}$ with average member set size greater than or equal to half of the size of it's universe $U(\mathcal{F})$.

Observation 1.3.3: *If the following inequality holds for a union-closed family $\mathcal{F}$, then $\mathcal{F}$ satisfies Conjecture 1.1.1.*

$$\frac{1}{|\mathcal{F}|} \cdot \sum_{A \in \mathcal{F}} |A| \geq \frac{1}{2}\ |U(\mathcal{F})|$$

Proof: Let m denote $|\mathcal{F}|$ and let n denote $|U(\mathcal{F})|$. Average size of member sets in $\mathcal{F}$ is greater than or equal to $\frac{n}{2}$, so the number of occurrences of all elements of $U(\mathcal{F})$ in $\mathcal{F}$ is at least $m \cdot \frac{n}{2}$. There are n elements in $|U(\mathcal{F})|$ and from the pigeonhole principle we conclude that at least one element must occur in at least $\frac{m}{2}$ member sets. $\square$

1.4 Equivalent formulations of the Union-closed sets conjecture

In this section, we discuss three equivalent formulations of the Union-closed sets conjecture. More details can be found in [4]. The union-closure property of families of sets has a dual intersection-closure property. Considering this, Conjecture 1.1.1 can be naturally reformulated in terms of these intersection-closed families of sets.

Conjecture 1.4.1: *Any finite intersection-closed family of at least two sets has an element that belongs to at most half of the member-sets.*

Observation 1.4.2: *Conjecture 1.4.1 is equivalent to Conjecture 1.1.1.*

Proof: To show the equivalence we first prove that a family is union-closed if and only if it's complement is intersection-closed. Let $\mathcal{F}$ be a family with universe $U(\mathcal{F})$, we define the complement of $\mathcal{F}$ as $\mathcal{F}^c = \{U(\mathcal{F}) \setminus A \ : \ A \in \mathcal{F}\}$. Let A, B be any two sets in $\mathcal{F}$, so clearly $U(\mathcal{F}) \setminus A$ and $U(\mathcal{F}) \setminus B$ are both in $\mathcal{F}^c$. Since $\mathcal{F}$ is union-closed, $(A \cup B) \in \mathcal{F}$ and thus $U(\mathcal{F}) \setminus (A \cup B)$ belongs to $\mathcal{F}^c$ And this is exactly the intersection $(U(\mathcal{F}) \setminus A) \cap (U(\mathcal{F}) \setminus B)$.

Next, we show that if $\mathcal{F}$ has an abundant element, then $\mathcal{F}^c$ has a rare element. Suppose that $\mathcal{F}$ has an abundant element a, which belongs to at least half of the member sets of $\mathcal{F}$. For every $A \in \mathcal{F}$ it holds that $a \in A \iff a \notin A^c$, so a belongs to at most half of the sets in $\mathcal{F}^c$ and is rare in $\mathcal{F}^c$. $\square$

An example of union-closed family and it's intersection-closed complementary family is shown in Figure 5.

$$
\begin{array}{cc}
\{1,2,3,4\} & \{1,2,3,4\} \\
\{1,2,3\},\ \{2,3,4\},\ \{1,2,4\} & \{1,3,4\},\ \{2,3,4\} \\
\{1,2\},\ \{1,3\},\ \{2,3\} & \{1,4\},\ \{2,4\},\ \{3,4\} \\
\{1\},\ \{2\} & \{1\},\ \{3\}\ \{4\} \\
\emptyset & \emptyset
\end{array}
$$

Figure 5: Example of union-closed family and it's intersection-closed complementary family.

Another equivalent formulation of Conjecture 1.1.1 is in terms of finite lattices. Using lattices to formulate a conjecture about union-closed families of sets is quite natural, since every union or intersection closed family forms a lattice [4]. We say an element $x \in L$ is join-irreducible if $\forall a, b \in L \ : \ x = a \vee b$ implies $x = a \ \vee \ x = b$.

Conjecture 1.4.3: *Let $(L, \preccurlyeq)$ be a finite lattice with at least two elements. Then there is a join-irreducible element $a \in L$ such that at most half of the elements of L are greater than or equal to a.*

To show the equivalence between this formulation and the Union-closed sets conjecture, it is sufficient to show that the lattice formulation is equivalent to Conjecture 1.4.1 as done in [4].

Theorem 1.4.4: *The lattice formulation 1.4.3 is equivalent to Conjecture 1.1.1.*

The Union-closed sets conjecture can also be formulated in terms of maximal stable sets of bipartite graphs. A *stable set* of a graph G is a subset of $V(G)$ where no two vertices are adjacent in G. A stable set is called *maximal* if no other vertex in G can be added to it.

Conjecture 1.4.5: *Let G be a bipartite graph with at least one edge. Then each bipartition class of G contains a vertex that belongs to at most half of the maximal stable sets of G.*

This formulation of the Union-closed sets conjecture was first stated by Bruhn, Charbit, Schaudt and Telle [5], who proved it's equivalence to Conjecture 1.4.1.

Theorem 1.4.6: *The graph formulation 1.4.5 is equivalent to Conjecture 1.1.1.*

The lattice and graph formulations resulted in a number of partial results of the Union-closed sets conjecture. For the lattice formulation the following result is known. Recall that a lattice $(L, \preccurlyeq)$ is lower semimodular if the implication $b \prec a \vee b \implies a \wedge b \prec a$ holds for all $a, b \in L$.

Theorem 1.4.7 (Reinhold [22]): *Lower semimodular lattices satisfy Conjecture 1.4.3.*

For the graph formulation the following two results were shown by Bruhn, Charbit, Schaudt, and Telle [5]. Recall that a graph is chordal if every cycle of length at least six has a chord, an edge connecting two vertices in the cycle that are distance greater than one apart in the cycle. Also recall that a graph is subcubic if all it's vertices have degrees less than or equal to three.

Theorem 1.4.8: *Chordal bipartite graphs satisfy Conjecture 1.4.5.*

Theorem 1.4.9: *Every subcubic bipartite graph satisfies Conjecture 1.4.5.*

1.5 Constant lower bound breakthrough

In 2022, Gilmer [14] made a breakthrough by proving a constant lower bound for the Union-closed sets conjecture.

Theorem 1.5.1: *Any finite union-closed family of sets $\mathcal{F} \neq \{\emptyset\}$ has an element that belongs to at least* 1% *of the member sets of* $\mathcal{F}$.

In his work, Gilmer claimed, that his results could be improved to $\frac{3-\sqrt{5}}{2} \approx 38\%$. Quickly after that Alweiss, Huang and Selke [1], Chase and Lovett [8], Sawin [25] and Pebody [19] independently proved it, greatly increasing the lower bound.

Theorem 1.5.2: *Any finite union-closed family of sets $\mathcal{F} \neq \{\emptyset\}$ has an element that belongs to at least* $\frac{3-\sqrt{5}}{2} \approx 38\%$ *of the member sets of* $\mathcal{F}$.

Later Cambie [6] showed an upper bound for the approach of Sawin [25]. This bound is smaller than one would hope for, approximately 0.038% greater than $\frac{3-\sqrt{5}}{2}$, which is still far from Conjecture 1.1.1.

Theorem 1.5.3: *Any finite union-closed family of sets $\mathcal{F} \neq \{\emptyset\}$ has an element that belongs to at least* $\frac{3-\sqrt{5}}{2} + 0.037952211\%$ *of the member sets of* $\mathcal{F}$.

Chase and Lovett [8] extended Theorem 1.5.2 to a generalization of Conjecture 1.1.1 using approximate union-closed families. For $0 \leq c \leq 1$ a c-approximate union-closed family $\mathcal{F}$ is a family if for at least a c-fraction of the pairs $A, B \in \mathcal{F}$ their union $A \cup B$ belongs to $\mathcal{F}$.

Theorem 1.5.4: *Let $\mathcal{F}$ be an $(1-\varepsilon)$-approximate union-closed set system, where $\varepsilon < \frac{1}{2}$, then there is an element which belongs to a $\frac{3-\sqrt{5}}{2} - \delta$ fraction of the sets in $\mathcal{F}$, where $\delta = 2\varepsilon\left(1 + \frac{\log(\frac{1}{\varepsilon})}{\log(|\mathcal{F}|)}\right)$.*

In their paper Chase and Lovett provide an example of an approximate union-closed family where the bound of $\frac{3-\sqrt{5}}{2}$ is optimal, see Example 1.4. in [8].

Before Gilmer's result, the best known lower bound on the frequency of the most common element was a simple observation by Knill [16].

Theorem 1.5.5: *Any union-closed family $\mathcal{F}$ on m member-sets has an element belonging to least $\frac{m-1}{\log_2(m)}$ of the member sets.*

For the sake of completeness, we include Knill's short proof, here explained with more details.

Proof: Choose a subset of the universe $S \subseteq U(\mathcal{F})$ minimal such that every non-empty member set of $\mathcal{F}$ intersects S. We show that for every $x \in S$ there is $A \in \mathcal{F}$ such that $A \cap S = \{x\}$. Otherwise $S \setminus x$ would have non-empty intersection with every $A \in \mathcal{F}$, which is a contradiction with the minimality of S. We extend this observation to any subset of S as follows. Consider the case for a 2-subset of S. We want to show that $\forall \{x, y\} \subseteq S$ there exists $A \in \mathcal{F}$ such that $A \cap S = \{x, y\}$. We know that there is $B \in \mathcal{F}$ such that $B \cap S = \{x\}$ and also there exists $C \in \mathcal{F}$ such that $C \cap S = \{y\}$. Since $\mathcal{F}$ is union-closed, we have $(B \cup C) \in \mathcal{F}$ and $(B \cup C) \cap S = \{x, y\}$. By a similar argument, each subset of S is an intersection $S \cap A$ for some $A \in \mathcal{F}$. From this we conclude that there must be at least as many member sets in $\mathcal{F}$ as there are subsets of S, $m \geq 2^{|S|}$ and so $\log_2(m) \geq |S|$. Clearly, there are at least $m - 1$ non-empty member-set of $\mathcal{F}$. We use that every non-empty member-set of $\mathcal{F}$ intersects S and we consider an element e of S most frequently appearing in there intersections, by the pidgeon-hole principle e appears in at least $\frac{m-1}{|S|}$ of these intersections. Thus, e belongs to at least $\frac{m-1}{\log_2(m)}$ member sets of $\mathcal{F}$. □

This result was later improved by Wójcik [29] to $\frac{2.4m}{\log_2(m)}$ for large m.

1.6 Frankl-complete families

The Union-closed sets conjecture can be attacked from many direction. Already in Section 1.3 we show that any union-closed family $\mathcal{F}$ containing a set of size 1 or 2 has an element that belongs to at least half of the member sets. This section talks about an extension of this idea, Frankl-complete families, which are small families for which any finite union-closed super-family satisfies Conjecture 1.1.1.

Definition 1.6.1 (Frankl-complete family): *A union-closed family of sets $\mathcal{F}_c$ is Frankl-complete (FC) if for any finite union-closed family $\mathcal{F} \supseteq \mathcal{F}_c$ an element of $U(\mathcal{F}_c)$ appears in at least half of the member sets of $\mathcal{F}$.*

Poonen gave necessary and sufficient conditions for a union-closed family to be Frankl-complete, see Theorem 1 in [20]. The author also proved that three 3-subsets of a 4-set is a Frankl-complete family.

Theorem 1.6.2: *The Union-closed sets conjecture holds for union-closed families containing three 3-subsets of a 4-set.*

Vaughan [27] extended Poonen's results and proved, among other results, that all five 4-subsets of a 5-set, or any ten 4-subsets of a 6-set generate an FC-family.

Theorem 1.6.3: *The Union-closed sets conjecture holds for union-closed families $\mathcal{F}$ satisfying one of the following.*

- *$\mathcal{F}$ contains all five 4-subsets of a 5-set*
- *$\mathcal{F}$ contains ten 4-subsets of a 6-set*

In 2007 Morris [21] shows a complete characterization of all Frankl-complete families on 5 elements. Additionally, he shows more sub-families that generate an FC-family. In the following theorem we mention three such sub-families, refer to [21] for more.

Theorem 1.6.4: *The Union-closed sets conjecture holds for every finite union-closed family $\mathcal{F}$ satisfying at least one of the following.*

- *$\mathcal{F}$ contains at least three 3-subsets of a 5-set.*
- *$\mathcal{F}$ contains at least four 3-subsets of a 6-set.*
- *$\mathcal{F}$ contains at least eight 4-subsets of a 6-set.*

Recently a full classification of all Frankl-complete families on at most 6 elements was given by Marić, Vučković and Živković [17].

1.7 Small families

A different direction to attack the Union-closed sets conjecture is to consider a restriction of the conjecture, one where the size of the biggest set is bounded by a given constant.

Conjecture 1.7.1: *Any finite union-closed family of sets $\mathcal{F} \neq \{\emptyset\}$ on at most n elements has an element that belongs to at least half of the member sets of $\mathcal{F}$ for a given bound n.*

To prove Conjecture 1.7.1 even for small n is a difficult task. We are able to check all families with brute-force only for the first few n. Indeed, as the power set of the universe of size n has 2^n elements, then the number of all possible families on n elements is 2^{2^n}. This sequence grows incredibly fast with n. Even for $n = 10$ the number of families on 10 elements is approximately $1.8 \cdot 10^{308}$, which is simply too much to iterate over.

Starting with Poonen, there have been many new ideas to push the bound in Conjecture 1.7.1 ever higher. In 1992, Poonen [20] proved that the Union-closed sets conjecture holds when the size of the biggest set in $\mathcal{F}$ is at most 7. Lo Faro [11] improved this bound to 9, followed by Marković [18], Bošnjak and Marković [3] and Vučković and Živković [28] who improved the bound to 10, 11 and 12 respectively.

Theorem 1.7.2: *The Union-closed sets conjecture holds for all families on at most 12 elements.*

The results of Vučković and Živković are the best known bound to date. In Chapter 2 we present our approach and improve the bound to 14.

A different result is the following theorem by Lo Faro [12] which provides a bound on the number of member sets based on the size of the biggest set.

Theorem 1.7.3: *Under the assumption that the union-closed sets conjecture fails, let m denote the minimum cardinality of $|U(\mathcal{F})|$ taken over all counterexamples $\mathcal{F}$ to the union-closed sets conjecture. Then any counterexample has at least $4m - 1$ member-sets.*

Combining Theorem 1.7.3 with Theorem 1.7.2 we get that the Union-closed sets conjecture holds for all families containing at most 50 member sets.

Corollary 1.7.3.1: *The Union-closed sets conjecture holds for all families with at most 50 member sets.*

1.8 Linear programming

In this section we recall the concept of linear programming and briefly demonstrate it's applications and geometric interpretation. For a detailed introduction to the topic we refer the reader, for instance, to the works of Chvátal [9] and Cook et al. [10].

Linear programming (LP) is a classical problem in mathematical optimization. The applications of linear programming range from solving practical problems like supply chain management, production planning and portfolio optimization to solving problems in numerous areas of theoretical mathematics.

The definition on linear programming is as follows.

Definition 1.8.1 (Linear programming): *Let $\boldsymbol{x}$ be a vector of n real variables, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{c} \in \mathbb{R}^n$. Minimizing the function $\boldsymbol{c}^T\boldsymbol{x}$ subject to constraints $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ written as*

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \boldsymbol{c}^T\boldsymbol{x}$$
$$\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$$

is called a linear programming problem.

The function $\boldsymbol{c}^T\boldsymbol{x}$ is called the *objective function.* A vector $\boldsymbol{x}$ satisfying the constraints $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ is called a *feasible solution.* A feasible solution which minimizes the objective function is called an *optimal solution.* Note that the objective function $\boldsymbol{c}^T\boldsymbol{x}$ is another way of writing $c_1x_1 + c_2x_2 + \ldots + c_nx_n$ which is a linear combination of elements of $\boldsymbol{x}$. An example objective function is plotted in Figure 6. Similarly, constraints $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ can be rewritten as a system of linear inequalities as follows.

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n &\geq b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n &\geq b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \ldots + a_{m,n}x_n &\geq b_m \end{aligned} \tag{1}$$

Each of these constraints defines a half-space. The intersection of all such half-spaces is a convex polytope. Each point of the polytope corresponds to a feasible solution and vice versa. In this way the polytope encodes all possible solutions of the problem and is called the *feasible region.*

One of the standard methods for solving linear programming problems is the simplex algorithm. It is established that if an LP problem has an optimal solution, then at least one optimal solution is a vertex of the feasible region. This is a consequence of the linearity of the objective function and of each constraint. The simplex algorithm starts in an initial vertex. There are multiple algorithms for choosing the initial vertex. In each iteration the algorithm considers adjacent vertices of the current vertex on the feasible region. It then compares the values of the objective function at these vertices. If the objective function at the current vertex is smaller than in all adjacent vertices, then the current vertex is an optimal solution. Otherwise the algorithm moves into the adjacent vertex with the lowest value of the objective function and continues to the next iteration.

We further explain linear programming on a simple example.

Example 1.8.2: *Suppose that we have a vector of two variables* $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$, *a vector* $\mathbf{c} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ *and want to minimize the objective function* $\mathbf{c}^T\boldsymbol{x} = x_1 + x_2$ *subject to constraints* $\boldsymbol{Ax} \geq \boldsymbol{b}$ *where*

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 \\ 1 & 1.4 \\ -1 & 2 \\ -3 & 1 \\ -1 & -1 \\ 0.1 & -1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} -1 \\ 3.5 \\ -1 \\ -15 \\ -12 \\ -6.5 \end{bmatrix}$$

The feasible region and the objective function are shown in Figure 6. Constraints $\boldsymbol{Ax} \geq \boldsymbol{b}$ can be written as follows.

$$\begin{aligned} 2x_1 - 1x_2 &\geq -1 \\ 1x_1 + 1.4x_2 &\geq 3.5 \\ -1x_1 + 2x_2 &\geq -1 \\ -3x_1 + 1x_2 &\geq -15 \\ -1x_1 - 1x_2 &\geq -12 \\ 0.1x_1 - 1x_2 &\geq -6.5 \end{aligned}$$

To solve this problem we use the simplex algorithm. We choose the vertex $\boldsymbol{x}_0 = \begin{bmatrix} 6.75 & 5.25 \end{bmatrix}^T$ as a starting point. Iterations of the simplex algorithm are shown in Table 1 and in Figure 7.

current vertex $\boldsymbol{x}_i$	$\mathbf{c}^T\boldsymbol{x}_i$	*adjacent vertices* $\boldsymbol{x}_i^a, \boldsymbol{x}_i^b$	$\mathbf{c}^T\boldsymbol{x}_i^a$	$\mathbf{c}^T\boldsymbol{x}_i^b$
$\boldsymbol{x}_0 = \begin{bmatrix} 6.75 & 5.25 \end{bmatrix}^T$	12	$\boldsymbol{x}_0^a = \begin{bmatrix} 5.8 & 2.4 \end{bmatrix}^T$, $\boldsymbol{x}_0^b = \begin{bmatrix} 5 & 7 \end{bmatrix}^T$	8.2	12
$\boldsymbol{x}_1 = \begin{bmatrix} 5.8 & 2.4 \end{bmatrix}^T$	8.2	$\boldsymbol{x}_1^a = \begin{bmatrix} 2.47 & 0.74 \end{bmatrix}^T$, $\boldsymbol{x}_1^b = \begin{bmatrix} 6.75 & 5.25 \end{bmatrix}^T$	3.21	12
$\boldsymbol{x}_2 = \begin{bmatrix} 2.47 & 0.74 \end{bmatrix}^T$	3.21	$\boldsymbol{x}_2^a = \begin{bmatrix} 0.55 & 2.1 \end{bmatrix}^T$, $\boldsymbol{x}_2^b = \begin{bmatrix} 5.8 & 2.4 \end{bmatrix}^T$	2.62	8.2
$\boldsymbol{x}_3 = \begin{bmatrix} 0.55 & 2.1 \end{bmatrix}^T$	2.62	$\boldsymbol{x}_3^a = \begin{bmatrix} 2.47 & 0.74 \end{bmatrix}^T$, $\boldsymbol{x}_3^b = \begin{bmatrix} 2.75 & 6.5 \end{bmatrix}^T$	3.21	9.25

Table 1: Iterations of the simplex algorithm

Solving linear programming problems with the simplex algorithm is generally fast and very efficient for practical applications, but there are instances in which the algorithm takes an exponential number iterations. There are polynomial algorithms for solving LP problems, for instance the ellipsoid algorithm. For practical applications the simplex algorithm is generally preferred.

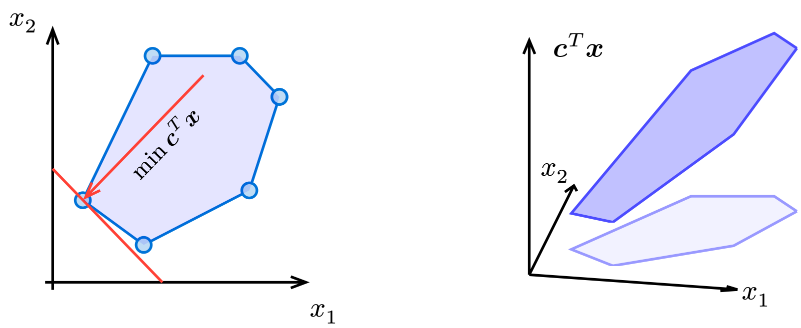


Figure 6: Geometric interpretation of example 1.8.2. On the left there is the feasible polygon, an arrow indicating the direction in which the objective function is minimized and a line perpendicular to the arrow indicating a contour of the objective function. On the right side is a plot of the objective function on the feasible polygon.

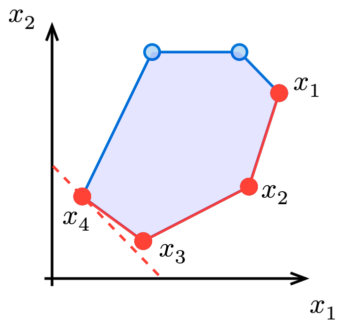


Figure 7: Steps of the simplex algorithm in example 1.8.2.

In a linear programming problem the variable $\boldsymbol{x}$ can take real values. However, there are scenarios in which only integer or binary values are allowed. For instance problems where the variable $\boldsymbol{x}$ describes an amount of people or whether a particular edge is present in a graph. In Section 2.1 we formulate a case of the Union-closed sets conjecture 1.1.1 as a linear programming problem with binary variables.

Definition 1.8.3 (Integer linear programming): *Let* $\boldsymbol{x}$ *be a vector of* n *integer variables,* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ *and* $\mathbf{c} \in \mathbb{R}^n$. *Minimizing the function* $\mathbf{c}^T\boldsymbol{x}$ *subject to constraints* $\boldsymbol{Ax} \geq \boldsymbol{b}$ *written as*

$$\min_{\boldsymbol{x} \in \mathbb{Z}^n} \boldsymbol{c}^T \boldsymbol{x}$$
$$\boldsymbol{Ax} \geq \boldsymbol{b}$$

is called an integer linear programming (ILP) problem.

For the sake of clarity we compare LP and ILP on an example. We consider an ILP problem with the objective function and constraints as in Example 1.8.2. The feasible region and the objective function of this ILP problem are shown in Figure 8. Observe that the integer restriction on each variable of the vector $\boldsymbol{x}$ makes the feasible region a discrete set of isolated points. We note that the feasible region of an ILP problem is a subset of the feasible region of an LP problem with the same constraints.

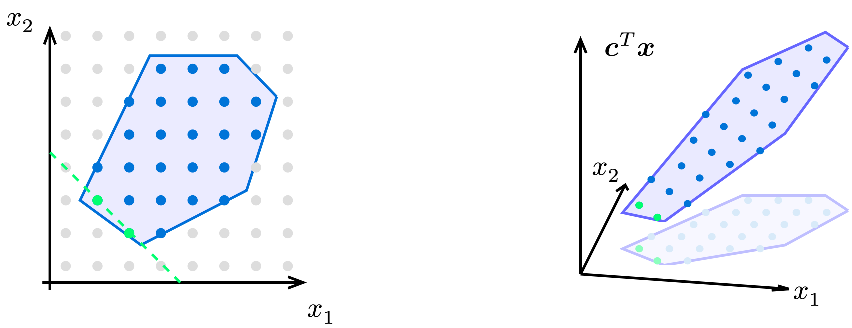


Figure 8: Example of an integer linear programming problem

Applying the simplex or the ellipsoid algorithms to integer linear programming problems is not possible. While there are algorithms for solving ILP problems, such as the branch and bound or the cutting plane method, these algorithms take an exponential number of iterations. In contrast to linear programming, ILP is known to be NP-Hard. Considering this, we may want to seek an approximate solution. An approximation we will use is called *relaxation.*

Definition 1.8.4 (Linear programming relaxation): *Given an integer linear programming problem* (2) *we define the relaxed linear programming problem to be* (3)*.*

$$\min_{\boldsymbol{x}\in\mathbb{Z}^n} \boldsymbol{c}^T\boldsymbol{x} \quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \tag{2}$$

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \boldsymbol{c}^T\boldsymbol{x} \quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \tag{3}$$

We note that the solution of the relaxed problem can be arbitrarily far from the solution of the original problem. Despite this, relaxation gives a bound on the objective function in an optimal solution of the original problem.

We established that the feasible region of an ILP problem is a subset of the feasible region of the corresponding relaxed problem. Using this and the fact that the objective function is a linear combination of elements of $\boldsymbol{x}$ we get that the inequality (4) holds for an optimal solution $\boldsymbol{x}_{LP}$ of the relaxed problem and an optimal solution $\boldsymbol{x}_{ILP}$ of the original ILP problem.

$$\boldsymbol{c}^T\boldsymbol{x}_{LP} \leq \boldsymbol{c}^T\boldsymbol{x}_{ILP} \tag{4}$$

We use this bound in Section 2.2 when formulating a restriction of the Union-closed sets conjecture as an LP problem.

Next, we turn our attention to duality in linear programming.

Definition 1.8.5 (Dual problem): *Consider an LP problem* (5)*. We call this problem primal. The dual problem to the primal problem is defined as* (6)*.*

Primal problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \boldsymbol{c}^T\boldsymbol{x} \quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \tag{5}$$

Dual problem

$$\max_{\boldsymbol{y}\in\mathbb{R}^m} \boldsymbol{b}^T\boldsymbol{y} \quad \boldsymbol{A}^T\boldsymbol{y} \leq \boldsymbol{c} \tag{6}$$

There are two fundamental theorems regarding duality in linear programming that describe the relationship between primal and dual solutions of LP problems.

Theorem 1.8.6 (Weak duality theorem): *For every LP problem* (5) *and it's dual* (6) *it holds: If* $\boldsymbol{x} \in \mathbb{R}^n$ *is a feasible solution of the primal and* $\boldsymbol{y} \in \mathbb{R}^m$ *is a feasible solution of the dual then* $\boldsymbol{b}^T\boldsymbol{y} \leq \boldsymbol{c}^T\boldsymbol{x}$*.*

Theorem 1.8.7 (Strong duality theorem): *If the primal* (5) *or the dual* (6) *has an optimal solution, then the other also has an optimal solution such that* $\boldsymbol{b}^T\boldsymbol{y} = \boldsymbol{c}^T\boldsymbol{x}$*.*

We use linear programming duality in Section 2.6 to quickly validate our results from Section 2.5 without having to compute the whole proof again.

2 Small families and linear programming

In this chapter we apply linear programming to a restriction of the Union-closed sets conjecture where the size of the universe is bounded. Our approach is to first formulate Conjecture 1.7.1 as an ILP problem in Section 2.1, which we then relax in Section 2.2. The rest of this chapter provides a detailed explanation of our approach to compensate for the relaxation, summarizes our computer assisted proof and discusses corollaries of the proof.

2.1 ILP-formulation of a restriction of the Union-closed sets conjecture

Recall the restriction of the Union-closed sets conjecture from Section 1.7. We bound the size of the universe by a given constant n. Also recall that the number of all subset of a set U is $2^{|U|}$, which is finite for any finite set U. The idea is to represent any family $\mathcal{F}$ on n elements by a collection of binary variables indicating whether a given subset of $\{1, ..., n\}$ is present in $\mathcal{F}$. In this way, we can represent any of the 2^{2^n} families on n elements as a binary vector of length 2^n. We illustrate this process on a simple example.

Example 2.1.1: *Let* $n = 3$ *and* $U = \{1, 2, 3\}$*. In each of the following cases we represent the family on the left as the binary vector of length* 2^3 *on the right.*

$$\begin{array}{c} \{1,2,3\} \\ \{1,2\},\{1,3\} \\ \{1\},\{2\} \\ \emptyset \end{array} \quad \longrightarrow \quad \begin{array}{cccccccc} x_\emptyset & x_{\{1\}} & x_{\{2\}} & x_{\{3\}} & x_{\{1,2\}} & x_{\{2,3\}} & x_{\{1,3\}} & x_{\{1,2,3\}} \\ [\,1, & 1, & 1, & 0, & 1, & 0, & 1, & 1\,] \end{array}$$

$$\begin{array}{c} \{1\},\{2,3\} \\ \emptyset \end{array} \quad \longrightarrow \quad \begin{array}{cccccccc} x_\emptyset & x_{\{1\}} & x_{\{2\}} & x_{\{3\}} & x_{\{1,2\}} & x_{\{2,3\}} & x_{\{1,3\}} & x_{\{1,2,3\}} \\ [\,1, & 1, & 0, & 0, & 0, & 1, & 0, & 0\,] \end{array}$$

Throughout the rest of this thesis, we denote x_A the variable indicating whether A is present in $\mathcal{F}$. Additionally, when working with a specific member set, for example $A = \{1, 2, 3\}$, we write x_{123} instead of $x_{\{1,2,3\}}$.

Currently we represent any family on n elements with binary variables, but only want to consider union-closed families. We express union-closure with the following set of constraints.

$$x_{A\cup B} + 1 \ \geq \ x_A + x_B \quad \forall\, A, B \in 2^U \tag{7}$$

Observation 2.1.2: *A family* $\mathcal{F}$ *is union-closed if and only if* $\mathcal{F}$ *satisfies the constraints (7)*

Proof: In the case that $A, B \notin \mathcal{F}$ we have $x_A = x_B = 0$ and thus $x_{A\cup B} + 1 \geq 0$ clearly holds. If exactly one of A, B belongs to $\mathcal{F}$, then $x_{A\cup B} + 1 \ \geq \ 1$ holds regardless of the value of $x_{A\cup B}$. In the case that both A and B belong to $\mathcal{F}$ we have $x_{A\cup B} + 1 \geq 2$, which holds if and only if $\mathcal{F}$ contains the union $A \cup B$. $\square$

We define a new variable a, which we call *abundance*. This variable is meant to represent the number of occurrences of the most frequent element. To this end, we add the following constraints for a.

$$a \geq \sum_{A \in 2^U :\, i \in A} x_A \qquad \forall i \in U \tag{8}$$

Simply put, for any element $i \in U$ abundance must be greater than or equal to the number of occurrences of i. For a minimal such that it satisfies constraints (8) the variable a is indeed equal to the number of occurrences of the most common element of the universe.

We would like to know whether every union-closed familiy $\mathcal{F}$ on n elements has an abundant element that belongs to at least half of the sets of $\mathcal{F}$. With $\mathcal{F}$ represented as a binary vector, we can express the number of member sets in $\mathcal{F}$ as a sum of all variables x_A. Thus a family $\mathcal{F}$ has an abundant element if and only if the following inequality holds for a minimal satisfying constraints (8).

$$a \geq \frac{1}{2} \sum_{A \in 2^U} x_A$$

In other words if $a - \frac{1}{2} \sum_{A \in 2^U} x_A$ is non-negative. Consequently if $a - \frac{1}{2} \sum_{A \in 2^U} x_A$ is non-negative for all families on at most n elements, then we know that Conjecture 1.7.1 holds.

$$a - \frac{1}{2} \sum_{A \in 2^U} x_A \tag{9}$$

We may equivalently ask whether the minimum of the expression $a - \frac{1}{2} \sum_{A \in 2^U} x_A$ over all families on at most n elements is non-negative. We note that the expression (9) is a linear combination of a and of all x_A and also that the constraints (7) and (8) are systems of linear inequalities. With this in mind we define the following ILP problem and prove it's equivalence to Conjecture 1.7.1.

Problem 2.1.3: *Let $U = \{1, ..., n\}$ for a given n. We define the following ILP problem.*

$$\begin{array}{lll}
\textit{Variables:} & x_A \in \{0,1\} & \forall\, A \in 2^U \\
& a \in \mathbb{Z} & \\
\textit{Minimize:} & a - \frac{1}{2} \sum\limits_{A \in 2^U} x_A & \\
\textit{Subject to:} & x_{A \cup B} + 1 \;\geq\; x_A + x_B & \forall\, A, B \in 2^U \\
& a \geq \sum\limits_{\substack{A \in 2^U \\ i \in A}} x_A & \forall\, i \in U
\end{array}$$

Observation 2.1.4: *Conjecture 1.7.1 is true for a given bound n if and only if an optimal solution of Problem 2.1.3 is non-negative for n.*

Proof: We showed in Observation 2.1.2 that constraints (7) are equivalent to a family $\mathcal{F}$ being union-closed, thus any feasible solution of Problem 2.1.3 is union-closed and all union-closed families on at most n elements are feasible. Constraints (8) and the fact that we are minimizing guarantee that in an optimal solution $\mathcal{F}$, the variable a represents the number of occurrences of the most frequent element in $\mathcal{F}$. If we subtract one half of the total number of member sets in $\mathcal{F}$ and get a non-negative answer, then $\mathcal{F}$ must have an abundant element. Because we were minimizing, any other feasible solution has a non-negative value of the objective function as well from which we conclude that every other union-closed family on at most n elements has an abundant element. Otherwise if an optimal solution $\mathcal{F}$ has negative objective value, then $\mathcal{F}$ has no abundant element and thus Conjecture 1.7.1 is false for n. $\square$

2.2 LP-relaxation

In the previous section we have formulated a restriction of the Union-closed sets conjecture with a bound on the size of the universe as an ILP Problem 2.1.3. However, solving this problem is very slow in practice. Recall from Section 1.8 that integer linear programming is an NP-hard problem. One way to deal with this is to relax the problem to an LP problem with real variables.

Problem 2.2.1: *Let $U = \{1, ..., n\}$ for a given n. We define an LP problem as follows.*

$$
\begin{array}{lll}
\textit{Variables:} & x_A \in [0,1] & \forall\, A \in 2^U \\
 & a \in \mathbb{R} & \\
\textit{Minimize:} & a - \frac{1}{2} \sum\limits_{A \in 2^U} x_A & \\
\textit{Subject to:} & x_{A \cup B} + 1 \;\geq\; x_A + x_B & \forall\, A, B \in 2^U \\
 & a \geq \sum\limits_{\substack{A \in 2^U \\ i \in A}} x_A & \forall\, i \in U
\end{array}
$$

Recall from Section 1.8 the inequality (4), which states that the objective value in an optimal solution of an ILP problem is greater than or equal to the objective value in an optimal solution of the corresponding relaxed problem. Thus if the optimal solution of the relaxed problem is non-negative, the optimal solution of the original ILP problem is also non-negative and therefore Conjecture 1.7.1 holds for n.

Observation 2.2.2: *For a given constant n it holds that if an optimal solution of Problem 2.2.1 will have non-negative objective value, then Conjecture 1.7.1 holds for n.*

Since Problem 2.2.1 is an LP problem, we can solve it quickly with an LP solver, we use Gurobi [15] as our solver of choice. In our experiments it turned out that the optimal solution of Problem 2.2.1 has negative objective value. We note that this does not mean that Conjecture 1.7.1 were false for a given n, because the reverse implication in Observation 2.2.2 does not hold. If the objective value in an optimal solution of Problem 2.2.1 is negative, then we can make no conclusions regarding Conjecture 1.7.1. In the next section we compensate the negativity by adding stronger constraints and splitting the proof into nultiple cases.

2.3 Proof architecture

In this section we try to improve the negative objective value in an optimal solution of Problem 2.2.1 by adding stronger constraints and splitting the proof into multiple cases.

Definition 2.3.1: *Let $\mathcal{F}$ be a family on at most n elements. We define a case containing $\mathcal{F}$ to be Problem 2.2.1 with member sets of $\mathcal{F}$ fixed, expressed by additional constraints (10).*

$$x_A = 1 \quad \forall A \in \mathcal{F} \tag{10}$$

While an optimal solution of Problem 2.2.1 has negative objective value, it turns out that cases containing large families ($|\mathcal{F}| > 20$) often have a non-negative optimal solution. From cases that have a non-negative optimal solution we are able to derive additional constraints for the rest of our proof.

Observation 2.3.2: *Let $\mathcal{F}$ be a family on at most n elements and suppose that the optimal solution of a case containing $\mathcal{F}$ has non-negative objective value. Now let $\mathcal{G}$ be a different family on at most n elements. In the case containing $\mathcal{G}$ we may assume that there are at most $|\mathcal{F}| - 1$ member sets of $\mathcal{F}$.*

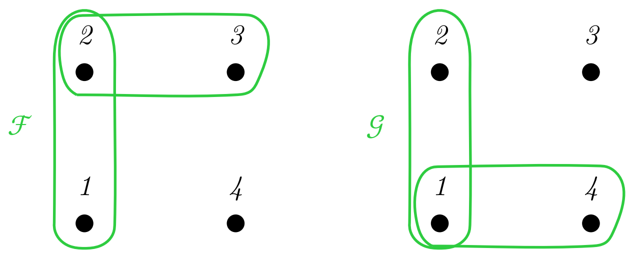
$$\sum_{A \in \mathcal{F}} x_A \ \leq \ |\mathcal{F}| - 1$$

Proof: Since the case containing $\mathcal{F}$ has non-negative optimal solution we can assume that there are at most $|\mathcal{F}| - 1$ member sets of $\mathcal{F}$ present. Otherwise the feasible region of case containing $\mathcal{G}$ would be a subset of case containing $\mathcal{F}$ and therefore have a non-negative optimal solution. $\square$

We generalize this observation to yield stronger constraints, but first we need to define what it means for two families to be *isomorphic*.

Definition 2.3.3: *Let $\mathcal{F}, \mathcal{G}$ be families of sets. We say that $\mathcal{F}$ and $\mathcal{G}$ are isomorphic, denoting $\mathcal{F} \cong \mathcal{G}$, if there exists a bijection $f : U(\mathcal{F}) \longrightarrow U(\mathcal{G})$ such that for all $A \in 2^{U(\mathcal{F})}$ it holds that $A \in \mathcal{F} \iff f(A) \in \mathcal{G}$, where $f(A)$ denotes $\{f(a) : a \in A\}$. We say that f is an isomorphism between $\mathcal{F}$ and $\mathcal{G}$.*

Example 2.3.4: *Let $\mathcal{F} = \{\{1, 2\}, \{2, 3\}\}$, $\mathcal{G} = \{\{1, 2\}, \{1, 4\}\}$ and $\mathcal{H} = \{\{1, 2\}, \{3, 4\}\}$.*



Clearly $\mathcal{F}$ is isomorphic to $\mathcal{G}$, an isomorphism between $\mathcal{F}$ and $\mathcal{G}$ is for example $f = \begin{cases} 2 \to 1 \\ 1 \to 4 \\ 3 \to 2 \\ 4 \to 3 \end{cases}$. On the other hand, $\mathcal{F} \ncong \mathcal{H}$ as $\mathcal{F}$ and $\mathcal{H}$ have different structure.

We note that isomorphism preserves the frequencies of all elements.

Observation 2.3.5: *Let $\mathcal{F}, \mathcal{G}$ be finite union-closed families. Let $f : U(\mathcal{F}) \to U(\mathcal{G})$ be an isomorphism between $\mathcal{F}$ and $\mathcal{G}$. Then for any element $e \in U(\mathcal{F})$ it holds that e is abundant in $\mathcal{F}$ if and only if the corresponding $f(e)$ is abundant in $\mathcal{G}$.*

For our purposes, when testing whether all finite union-closed families on at most n elements have an abundant element, we can think of isomorphic families as being equivalent in terms of abundance.

Lemma 2.3.6: *Let $\mathcal{F}$ and $\mathcal{G}$ be families on at most n elements and let the case containing $\mathcal{F}$ have a non-negative optimal solution. If $\mathcal{G}$ contains a sub-family $\mathcal{G}'$ such that $\mathcal{G}' \cong \mathcal{F}$ then we can skip the case containing $\mathcal{G}$.*

Proof: Since a case containing $\mathcal{F}$ has a non-negative solution, we know that all super-families of $\mathcal{F}$ have an abundant element. In the case that $\mathcal{G}' = \mathcal{G}$ and therefore $\mathcal{G} \cong \mathcal{F}$ we use Observation 2.3.5 and get that all super-families of $\mathcal{G}$ have an abundant element. Otherwise let $\mathcal{G}'$ be a proper sub-family of $\mathcal{G}$. Clearly, every super-family of $\mathcal{G}$ is also a super-family of $\mathcal{G}'$. Since $\mathcal{G}' \cong \mathcal{F}$ we use Observation 2.3.5 and get that all super-families of $\mathcal{G}'$, including all super-families of $\mathcal{G}$, have an abundant element, which is what we wanted. □

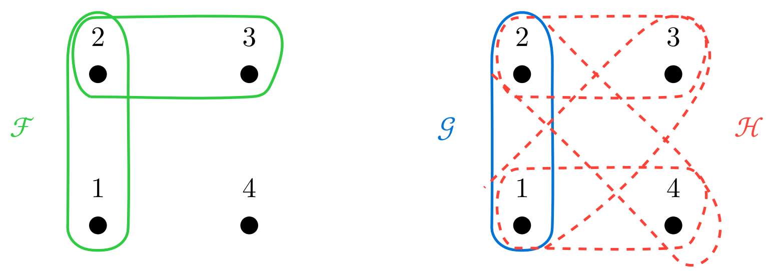
For two given families $\mathcal{F}, \mathcal{G}$ checking whether $\mathcal{G}$ contains a subgrap isomorphic to $\mathcal{F}$ is not an easy task. Luckily, the SageMath [26] library provides a function for this task, which we use in our proof. With this in mind we prove the following lemma which we use to generate strong constraints from previous non-negative cases.

Lemma 2.3.7: *Let $\mathcal{F}, \mathcal{G}$ be families on at most n elements and let the case containing $\mathcal{F}$ have a non-negative optimal solution. In the case containing $\mathcal{G}$ we may assign zero to all variables x_A for which $\{A\} \cup \mathcal{G}$ is isomorphic to a super-family of $\mathcal{F}$, resulting in the following set of constraints.*

$$x_A = 0 \qquad \forall A \in (2^U \setminus \mathcal{G}) \;:\; (\{A\} \cup \mathcal{G}) \cong \mathcal{F}' \supseteq \mathcal{F}$$

Proof: Let $A \in 2^U$ be a set such that $\{A\} \cup \mathcal{G}$ is isomorphic to $\mathcal{F}' \supseteq \mathcal{F}$ and let a case containing $\mathcal{F}$ has a non-negative optimal solution. If A is assigned one, we use Lemma 2.3.6 and get that we can skip this case as all super-families of $\{A\} \cup \mathcal{G}$ have an abundant element. As a result we assign zero to the variable x_A in the case containing $\mathcal{G}$. □

Example 2.3.8: *Let $U = \{1, 2, 3, 4\}$ and $\mathcal{F} = \{\{1, 2\}, \{2, 3\}\}$ and suppose that the case containing $\mathcal{F}$ has a non-negative optimal solution. Now consider the family $\mathcal{G} = \{\{1, 2\}\}$. In the case containing $\mathcal{G}$ we may, as a corollary of Lemma 2.3.7, assume that each set of $\mathcal{H} := \{\{2, 3\}, \{1, 3\}, \{1, 4\}, \{2, 4\}\}$ can be assigned zero.*



As a corollary we keep a collection $\mathcal{C}$ containing all families $\mathcal{F}$ for which a case containing $\mathcal{F}$ had a non-negative optimal solution. In each subsequent case containing another family $\mathcal{G}$ we first check if $\mathcal{G}$ contains a sub-family $\mathcal{G}'$ such that $\mathcal{G}' \cong \mathcal{F}$ for some $\mathcal{F} \in \mathcal{C}$. If it does, then we skip the case containing $\mathcal{G}$. Otherwise we we assign zero to all variables x_A for which $\{A\} \cup \mathcal{G}$ is isomorphic to a super-family of at least one $\mathcal{F} \in \mathcal{C}$.

Our approach is to choose cases systematically. We pick a specific class of families and run a case containing $\mathcal{F}$ for every non-isomorphic $\mathcal{F}$ in the given class. If all of them are non-negative, then we may add an additional constraint. These classes of families are k-uniform families with m member sets. A family $\mathcal{F}$ is *k-uniform* if every member set of $\mathcal{F}$ has size k.

Lemma 2.3.9: *Let $C_{k,m}$ be the class of all non-isomorphic k-uniform families with m member sets. Assume that all cases containing families $\mathcal{F} \in C_{k,m}$ have non-negative objective value. Then we may assume that there are at most $m-1$ sets of size k.*

$$\sum_{\substack{A \in 2^U \\ |A|=k}} x_A \le m-1 \tag{11}$$

Proof: Let $\mathcal{H}$ be a family on at most n elements containing m member sets of size k. And assume that all member sets $A \in \mathcal{H}$ have $x_A = 1$. Clearly $\mathcal{H}$ is a k-uniform family with m member sets and must be isomorphic to a family $\mathcal{F} \in C_{k,m}$. Since the case containing $\mathcal{F}$ has a non-negative solution, we can use Lemma 2.3.6 to conclude that all super-families of $\mathcal{H}$ have an abundant element. We therefore assume that at least one member set of $\mathcal{H}$ is missing. $\square$

We pick k, m and prove all cases containing non-isomorphic k-uniform families with m member sets. Then using Lemma 2.3.9 we add constraints (11) and pick a new pair of k, m. For a given k and a sufficiently large m we are usually able to prove all cases containing non-isomorphic k-uniform families with m member sets. We then add constraints (11) and try to gradually lower the number of member sets m, obtaining ever stronger constraints. The goal is to eventually prove that for a given k the sum of all k-sets is equal to zero.

Recall Observation 1.3.3 from Section 1.3 stating that if the average size of all member sets is greater than or equal to $\frac{n}{2}$ then there exists an abundant element. In our proof we guarantee this bound on average member set size by proving Lemma 2.3.9 for all k up to $\lceil \frac{n}{2} - 1 \rceil$ and for $m = 1$ yielding the following constraint.

$$\sum_{\substack{A \subseteq 2^U \\ |A| \le \lceil \frac{n}{2} - 1 \rceil}} x_A = 0 \tag{12}$$

This constraint guarantees that all member sets are of size at least $\frac{n}{2}$ and thus the average member set size is at least $\frac{n}{2}$ as well. Therefore by Observation 1.3.3 we conclude that all families satisfying this constraint have an abundant element, which we use to finish our proof.

Another class of families from which we want to derive constraints are k-dominated families with m member sets. A family of sets $\mathcal{F}$ is k-dominated if there exists a k-subset of U which belongs to all member sets of $\mathcal{F}$.

Lemma 2.3.10: *Let $C_{k,m}$ be the class of all non-isomorphic k-dominated families with m member sets. Assume that all cases containing families $\mathcal{F} \in C_{k,m}$ have non-negative* *objective value. Then we can assume that for all $X \subset U$ such that $|X| = k$ there are at most $m - 1$ supersets of X.*

$$\sum_{\substack{A \in 2^U \\ X \subseteq A}} x_A \leq m - 1 \qquad \forall X \subset U : |X| = k \tag{13}$$

Proof: Let $\mathcal{H}$ be a k-dominated family on at least m member sets. Since $\mathcal{H}$ is k-dominated, it contains a sub-family $\mathcal{H}'$ isomorphic to some $\mathcal{F} \in C_{k,m}$. Using Lemma 2.3.6 we get that all super-families of $\mathcal{H}$ are non-negative. Therefore we assume that there are at most $m - 1$ sets all containing a fixed subset of size k. □

In our experiments it turned out that there are families of sets, for which cases have particularly negative optimal solutions. This includes k-dominated families with few member sets. We deploy a strategy to deal with these families at the beginning of our proof.

Observation 2.3.11: *Let $\mathcal{F}$ be a family and let the case containing $\mathcal{F}$ be negative. We choose an arbitrary number of sets $A \in (2^U \setminus \mathcal{F})$ and add constraints $x_A = 0$ for all of the sets A such that the case containing $\mathcal{F}$ with the constraints $x_A = 0$ has a non-negative optimal solution. Thus if the constraints $x_A = 0$ were valid then the case containing $\mathcal{F}$ is proved. We can obtain these constraints using Lemma 2.3.7 for cases containing $\{A\} \cup \mathcal{F}$. If the optimal solution of cases containing $\{A\} \cup \mathcal{F}$ are non-negative for all $\{A\}$, then using Lemma 2.3.7 we obtain constraints $x_A = 0$ and the case containing $\mathcal{F}$ is proved. If the for some A the case containing $\{A\} \cup \mathcal{F}$ has a negative optimal solution, then we may use the same process on $\{A\} \cup \mathcal{F}$ recursively, proving another set of constraints $x_B = 0$ until we get that $\{A\} \cup \mathcal{F}$ has a non-negative optimal solution.*

This is a long recursive process, which we only really want to run for certain particularly negative cases. For example let $\mathcal{F}$ be a k-dominated family with few member sets, then solving the case containing $\mathcal{F}$ with this method allows us to skip a lot of cases using Lemma 2.3.6 because $\mathcal{F}$ has few member sets. Additionally, because $\mathcal{F}$ is k-dominated, we are a step closer to proving all cases containing non-isomorphic k-dominated families with m member sets and using Lemma 2.3.10 to obtain strong constraints for a small m.

Another direction to tackle the negativity of cases is by proving additional stronger constraints, which we present in the next section.

2.4 Strengthening constraints

Recall from Section 1.3 that if a finite union-closed family $\mathcal{F}$ contains a member set A of size one or two, then at least one of the elements of A is abundant in $\mathcal{F}$. Therefore in Problem 2.2.1 we may assume that any feasible solution contains no set of size one or two. We express this with the following constraints.

$$\sum_{\substack{A \in 2^U \\ |A| = 1}} x_A = 0 \qquad \sum_{\substack{B \in 2^U \\ |B| = 2}} x_B = 0 \tag{14}$$

Also recall from Section 1.6 the results of Poonen [20], Vaughan [27] and Morris [21] regarding Frankl-complete families. Since any finite union-closed family $\mathcal{F}$ containing an

FC-family has an abundant element, we can assume that no feasible solution contains an FC-family. Theorem 1.6.2 states that three 3-subsets of a 4-set form a FC-family. Thus as a corollary of Theorem 1.6.2 we may assume that there are at most two 3-subsets of any 4-set in $\mathcal{F}$.

$$\sum_{\substack{A \subseteq M \\ |A|=3}} x_A \leq 2 \qquad \forall M \subseteq U : |M| = 4 \tag{15}$$

Similar constraints can be constructed from Theorems 1.6.3 and 1.6.4 simply with different bounds and set sizes. We note that all such constraints are linear inequalities.

Next, we show that it is sufficient to only consider families on exactly n elements. To this end, we prove the following lemma.

Lemma 2.4.1: *Let C_i be the class of all union-closed families of sets whose largest set is of size precisely i, and let n and m be positive integers such that $n > m$. If every family from C_n has an abundant element, then every family from C_m has an abundant element.*

Proof: Consider a family $\mathcal{F}_m \in C_m$ and assume that every family $\mathcal{F}_n \in C_n$ has an abundant element. If all elements of $U(\mathcal{F}_m)$ are abundant in $\mathcal{F}_m$ then the implication clearly holds. Otherwise denote e the element of $U(\mathcal{F}_m)$ that belongs to less than half of the member sets of $\mathcal{F}_m$. We add $n - m$ new elements to all sets $A \in \mathcal{F}_m : e \in A$ and denote the result $\mathcal{F}'_m$. Clearly $\mathcal{F}'_m$ is union-closed and belongs to C_n. The number of member sets does not change by adding elements to them, so $|\mathcal{F}_m| = |\mathcal{F}'_m|$. Since e is not abundant in $\mathcal{F}_m$, e and all of the newly added elements are not abundant in $\mathcal{F}'_m$. On the other hand all of the elements abundant in $\mathcal{F}_m$ are also abundant in $\mathcal{F}'_m$. Combining this we get that $\mathcal{F}_m$ has an abundant element if and only if $\mathcal{F}'_m$ has an abundant element, and using the assumption that every family in C_n has an abundant element we conclude that $\mathcal{F}_m$ has an abundant element. $\square$

Corollary 2.4.1.1: *It is sufficient to check all families on exactly n elements, so we may assume that the universe $U = \{1, ..., n\}$ is present in $\mathcal{F}$, since every union-closed family contains it's universe, as shown in Section 1.3.*

$$x_U = 1 \tag{16}$$

Recall from Section 2.1 the union-closure constraints (7). In practice, the number of these constraints is high and having too many constraints slows down the LP solver by increasing the number of vertices of the feasible region. Recall that there are 2^n variables x_A in our program and we define constraints (7) for every pair of distinct variables x_A, yielding $\binom{2^n}{2}$ constraints in total. Instead, we present a different approach. Consider a case containing some family $\mathcal{F}$ and define the following set of constraints.

$$x_{A \cup B} + 1 \; \geq \; x_A + x_B \quad \forall A \in \mathcal{F} \;\; \forall B \in 2^U \tag{17}$$

Clearly, for each set $A \in \mathcal{F}$ the variable x_A is equal to 1, which simplifies the inequality.

$$x_{A \cup B} \; \geq \; x_B \quad \forall A \in \mathcal{F} \;\; \forall B \in 2^U \tag{18}$$

In our experience, in an optimal solution of most cases, variables associated with small member-sets tend to have high value, while variables associated with larger sets tend to have smaller value. Our aim is to, through constraints detailed in this section, guarantee that the value of smaller sets will propagate up to variables representing larger sets. We say we *boost* the larger sets. We note that each of the constraints (18) is a stronger constraint than any of the constraints (7). In the constraints (7) the sum on the right side needs to be greater than 1 in order to propagate value to the variable $x_{A \cup B}$. On the other hand, in (18) the value of x_B gets propagated to $x_{A \cup B}$, since there is no additive constant on the left side of the inequality.

We have successfully lowered the number of constraints (7) by using constraints (17) in our proof instead, and reduced the time the LP solver takes to solve each case. However, by omitting some of the constraints (7) we expand the feasible region and as a result lower the objective value of an optimal solution of each case. We introduce more types of constraints to improve this.

Choose a set of interest S that we want to boost, usually one of the larger sets in 2^U. We boost S by it's subsets of size $|S| - 1$ using the following inequality.

$$(|S| - 1) \cdot x_S + 1 \;\; \geq \sum_{\substack{B \subset S \\ |B| = |S| - 1}} x_B \tag{19}$$

For example when $S = \{1, 2, 3\}$ we get the following.

$$2 \cdot x_{123} + 1 \;\geq\; x_{12} + x_{13} + x_{23}$$

Clearly, when one of the subsets on the right hand side is present, then the inequality holds as the left hand side is at least one. When two of them are present then by union-closure x_{123} is present too and thus we get $3 \geq 2$. Then if all three of the sets on the left hand side are present the inequality still holds. We generalize this observation to the constraints (19).

Observation 2.4.2: *Every finite union-closed family satisfies constraints (19).*

Proof: Let $\mathcal{F}$ be a union-closed family and let $S \in 2^U$. If none of the $(|S| - 1)$-subsets of S belongs to $\mathcal{F}$, then the sum on the right hand side is equal to zero and the inequality clearly holds. In the case that exactly one $(|S| - 1)$-subset of S belongs to $\mathcal{F}$, then the left hand side is equal to one and the inequality still clearly holds. If two or more $(|S| - 1)$-subsets of S belong to $\mathcal{F}$, then the right side of the inequality is equal to $|S|$, because the union of any two $(|S| - 1)$-subsets of S is S which by union-closure belongs to $\mathcal{F}$. The number of $(|S| - 1)$-subsets of S is equal to $|S|$ and therefore the inequality holds. □

In a few special cases we can improve the coefficient $|S| - 1$ to a smaller number, obtaining a stronger inequality. We do this using FC-families. Recall for example Theorem 1.6.2 from Section 1.6. We combine the fact that a family $\mathcal{F}$ has an abundant element if there are 3 or more 3-subsets of a 4-set with constraints (19) and get the following inequality for $S = \{1, 2, 3, 4\}$.

$$1 \cdot x_{1234} + 1 \;\geq\; x_{123} + x_{124} + x_{134} +_{234}$$

For the last set of constraints, let $\mathcal{F}$ be a family of sets and let A be a member set of $\mathcal{F}$ or a union of member sets of $\mathcal{F}$. Now consider the case containing $\mathcal{F}$, clearly $x_A =$

1. Next, let $B \subset A$ and let $C \subset (U \setminus A)$. We construct the following inequality and prove it's validity.

$$x_{A \cup C} + (|B| - 1) \cdot x_{B \cup C} + 1 \;\geq\; x_A + \sum_{\substack{D \subset B \\ |D| = |B|-1}} x_{D \cup C} \tag{20}$$

Since $x_A = 1$ the inequality can be simplified to the following.

$$x_{A \cup C} + (|B| - 1) \cdot x_{B \cup C} \;\geq\; \sum_{\substack{D \subset B \\ |D| = |B|-1}} x_{D \cup C}$$

Which is a strong constraint because the left hand side lacks an additive constant. Thus the weight of any of the $D \cup C$ will propagate to the left hand side. We provide a simple example of constraints (20) for $A = \{0, 1, 2, 3\}$, $B = \{0, 1, 2\}$ and $C = \{5, 6\}$.

$$x_{012356} + 2 \cdot x_{01256} + 1 \;\geq\; x_{0123} + x_{0156} + x_{0256} + x_{1256}$$

One can easily check that this inequality holds. We know that $x_A = 1$. In the case $x_{0156} + x_{0256} + x_{1256} = 0$ the inequality clearly holds as the left hand side is at least one. When $x_{0156} + x_{0256} + x_{1256} = 1$, for example if $x_{0165} = 1$, then the inequality holds because the union $\{0, 1, 2, 3\} \cup \{0, 1, 5, 6\}$ is equal to $\{0, 1, 2, 3, 5, 6\}$, which is present on the left hand side and thus the inequality holds tight. In the case that $x_{0156} + x_{0256} + x_{1256} = 2$ the union of any two of the sets yields $\{0, 1, 2, 5, 6\}$, the left hand side is equal to four and the inequality holds. The inequality also holds if all variables on the right hand side are equal to one. We generalize this observation to constraints (20).

Observation 2.4.3: *Every finite union-closed family satisfies constraints (20).*

Proof: The union of any two $D \cup C$ always yields $B \cup C$ as D is $(|B| - 1)$-subset of B. Additionally, the union of A and any $D \cup C$ is equal to $A \cup C$ as $D \subset A$. Thus if there are at least two $D \cup C$ present, the inequality holds. The number of all $D \cup C$ is $|B|$, thus if all of $D \cup C$ are present we have $|B| + 1$ on the right hand side and $|B| + 1$ on the left hand side and the inequality holds tight. In the case that only one of $D \cup C$ is present, then by union-closure $A \cup C$ must also be present and the inequality holds. When none of the $D \cup C$ is present the inequality holds trivially, which completes the proof. $\square$

2.5 Outline of the proof

In this section we present an outline of our proof of Conjecture 1.7.1 for $n = 14$. First, we assume that there are no sets of size one and two and obtain a bound on the number of k-sets for $k = 3, 4, 5, 6$.

Description	*Constraint*	*Using*
Prove all cases containing non-isomorphic 3-uniform families with 6 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\| = 3}} x_A \;\leq\; 5$	Lemma 2.3.9

Recursively prove a case containing a 3-dominated 4-uniform family $\mathcal{F}$ with 6 member sets.	$\mathcal{F} \in \mathcal{C}, \quad \sum_{A \in \mathcal{F}} x_A \leq \lvert\mathcal{F}\rvert - 1$	Observation 2.3.11
Prove all cases containing non-isomorphic 2-dominated 4-uniform families with 8 member sets.	$\sum_{\substack{A \in 2^U \\ a,b \in A \\ \lvert A\rvert=4}} x_A \leq 7 \quad \forall a, b \in U$	Lemma 2.3.10
Prove all cases containing non-isomorphic 1-dominated 4-uniform families with 7 member sets.	$\sum_{\substack{A \in 2^U \\ a \in A \\ \lvert A\rvert=4}} x_A \leq 6 \quad \forall a \in U$	Lemma 2.3.10
Prove all cases containing non-isomorphic 4-uniform families with 6 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=4}} x_A \leq 5$	Lemma 2.3.9
Recursively prove a case containing the family of sets $\mathcal{F} = \{\{0,1,2,3,4\},$ $\{0,1,5,6,7\}$,$\{2,3,5,6,8\}\}$	$\mathcal{F} \in \mathcal{C}, \quad \sum_{A \in \mathcal{F}} x_A \leq \lvert\mathcal{F}\rvert - 1$	Observation 2.3.11
Recursively prove a case containing the family of sets $\mathcal{F} = \{\{0,1,2,3,4\},$ $\{0,1,2,5,6\}$,$\{0,3,5,7,8\}\}$	$\mathcal{F} \in \mathcal{C}, \quad \sum_{A \in \mathcal{F}} x_A \leq \lvert\mathcal{F}\rvert - 1$	Observation 2.3.11
Recursively prove a case containing a 4-dominated 5-uniform family with 6 member sets.	$\mathcal{F} \in \mathcal{C}, \quad \sum_{A \in \mathcal{F}} x_A \leq \lvert\mathcal{F}\rvert - 1$	Observation 2.3.11
Prove all cases containing non-isomorphic 3-dominated 5-uniform families with 12 member sets.	$\sum_{\substack{A \in 2^U \\ a,b,c \in A \\ \lvert A\rvert=5}} x_A \leq 11 \quad \forall a, b, c \in U$	Lemma 2.3.10
Prove all cases containing non-isomorphic 2-dominated 5-uniform families with 10 member sets.	$\sum_{\substack{A \in 2^U \\ a,b \in A \\ \lvert A\rvert=5}} x_A \leq 9 \quad \forall a, b \in U$	Lemma 2.3.10

Prove all cases containing non-isomorphic 1-dominated 5-uniform families with 9 member sets.	$\sum_{\substack{A \in 2^U \\ a \in A \\ \lvert A\rvert=5}} x_A \leq 8 \quad \forall a \in U$	Lemma 2.3.10
Prove all cases containing non-isomorphic 5-uniform families with 7 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=5}} x_A \leq 6$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 7 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=6}} x_A \leq 6$	Lemma 2.3.9

The rest of the proof is straightforward. We iteratively prove slightly stronger constraints until we get that no sets of size less that 7 are present.

Description	*Constraint*	*Using*
Prove all cases containing non-isomorphic 3-uniform families with 5 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=3}} x_A \leq 4$	Lemma 2.3.9
Prove all cases containing non-isomorphic 4-uniform families with 5 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=4}} x_A \leq 4$	Lemma 2.3.9
Prove all cases containing non-isomorphic 5-uniform families with 5 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=5}} x_A \leq 4$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 5 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=6}} x_A \leq 4$	Lemma 2.3.9
Prove all cases containing non-isomorphic 3-uniform families with 4 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=3}} x_A \leq 3$	Lemma 2.3.9
Prove all cases containing non-isomorphic 4-uniform families with 4 member sets.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=4}} x_A \leq 3$	Lemma 2.3.9

Prove all cases containing non-isomorphic 5-uniform families with 4 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=5}} x_A \leq 3$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 4 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=6}} x_A \leq 3$	Lemma 2.3.9
Prove all cases containing non-isomorphic 3-uniform families with 3 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=3}} x_A \leq 2$	Lemma 2.3.9
Prove all cases containing non-isomorphic 4-uniform families with 3 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=4}} x_A \leq 2$	Lemma 2.3.9
Prove all cases containing non-isomorphic 5-uniform families with 3 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=5}} x_A \leq 2$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 3 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=6}} x_A \leq 2$	Lemma 2.3.9
Prove all cases containing non-isomorphic 3-uniform families with 2 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=3}} x_A \leq 1$	Lemma 2.3.9
Prove all cases containing non-isomorphic 4-uniform families with 2 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=4}} x_A \leq 1$	Lemma 2.3.9
Prove all cases containing non-isomorphic 5-uniform families with 2 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=5}} x_A \leq 1$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 2 member sets.	$\sum_{\substack{A \in 2^U \\ \|A\|=6}} x_A \leq 1$	Lemma 2.3.9

Prove all cases containing non-isomorphic 3-uniform families with 1 member set.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=3}} x_A \leq 0$	Lemma 2.3.9
Prove all cases containing non-isomorphic 4-uniform families with 1 member set.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=4}} x_A \leq 0$	Lemma 2.3.9
Prove all cases containing non-isomorphic 5-uniform families with 1 member set.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=5}} x_A \leq 0$	Lemma 2.3.9
Prove all cases containing non-isomorphic 6-uniform families with 1 member set.	$\sum_{\substack{A \in 2^U \\ \lvert A\rvert=6}} x_A \leq 0$	Lemma 2.3.9

Now, the average size of the remaining sets is clearly greater than or equal 7, so we use Observation 1.3.3 to conclude that all finite Union-closed families on at most 14 elements have an abundant element.

2.6 Verification

We provide a way to verify our results from Section 2.5. When running our proof we save the dual solution of each case, along with the constraints of the case. From this we are able to verify each case as follows.

Let $\mathcal{F}$ be a family and consider the case containing $\mathcal{F}$. We save the matrix $\boldsymbol{A}$ together with the vector $\boldsymbol{b}$ and the dual solution $\boldsymbol{y}$ of the case containing $\mathcal{F}$ and show that we can verify the non-negativity of the case containing $\mathcal{F}$ without running the LP solver.

Lemma 2.6.1: *Let $\mathcal{F}$ be a family of sets and let a case containing $\mathcal{F}$ have an optimal solution $\boldsymbol{x}$. Let $\boldsymbol{y} \in \mathbb{R}^m$ be a vector of length m where m is the number of constraints in the case containing $\mathcal{F}$. Let $\boldsymbol{Ax} \geq \boldsymbol{b}$ denote the constraints in the case containing $\mathcal{F}$ and denote $\boldsymbol{c}$ the coefficients of the objective function. If $\boldsymbol{A}^T\boldsymbol{y} \leq \boldsymbol{c}$ and $\boldsymbol{b}^T\boldsymbol{y} \geq 0$ then the case containing $\mathcal{F}$ has a non-negative optimal solution.*

Proof: The primal problem of the case containing $\mathcal{F}$ is minimizing $\boldsymbol{c}^T\boldsymbol{x}'$ subject to constraints $\boldsymbol{Ax} \geq \boldsymbol{b}$. Recall from Section 1.8 that the dual problem the case containing $\mathcal{F}$ is maximizing $\boldsymbol{b}^T\boldsymbol{y}'$ subject to constraints $\boldsymbol{A}^T\boldsymbol{y}' \leq \boldsymbol{c}$. It is immediate that $\boldsymbol{A}^T\boldsymbol{y} \leq \boldsymbol{c}$ is testing feasibility of $\boldsymbol{y}$. Moreover, $\boldsymbol{b}^T\boldsymbol{y} \geq 0$ means that the objective value of $\boldsymbol{y}$ is non-negative in the dual problem. Now we can use the Weak duality theorem 1.8.6 from Section 1.8 and get that the objective value of any every feasible solution $\boldsymbol{x}'$ of the primal problem is at least $\boldsymbol{b}^T\boldsymbol{y}$ and thus $\boldsymbol{c}^T\boldsymbol{x}' \geq \boldsymbol{b}^T\boldsymbol{y} \geq 0$ for all feasible solutions $\boldsymbol{x}'$ and therefore the case containing $\mathcal{F}$ has a non-negative optimal solution. □

For every non-negative case we save $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{y}$ and are able to verify each case afterwards. To this end we provide a verification program, that verifies any case in our proof in accordance with Lemma 2.6.1. The verification program, data of each case and our proof framework will all be publicly available shortly after this thesis when we publish our paper on this topic.

2.7 Results

In Section 2.5 we provide an outline of a proof of Conjecture 1.7.1 for $n = 14$, which can be verified as discussed in Section 2.6.

Theorem 2.7.1: *The Union-closed sets conjecture holds for all finite union-closed families on at most 14 elements.*

Using Theorem 1.7.3 by Lo Faro [11] we get that the Union-closed sets conjecture holds for all families with at most 58 member sets.

Corollary 2.7.1.1: *The Union-closed sets conjecture holds for all finite union-closed families with at most 58 member sets.*

Recall from Section 1.4 the lattice formulations of the Union-closed sets conjecture. Our results in Theorem 2.7.1 can be interpreted in terms of lattices by modifying the proof of the equivalence of the lattice formulation to the Union-closed sets conjecture in [24] and in [4].

Observation 2.7.2: *If the Union-closed sets conjecture holds for all finite union-closed families on at most n elements, then the lattice formulation 1.4.3 holds for all finite lattices with at least two elements and at most 14 join-irreducible elements.*

Proof: Let $(L, \preccurlyeq)$ be a finite lattice with at least two elements and at most n join-irreducible elements and denote $S(x)$ the set of all join-irreducible elements of L that are less than or equal to x. Clearly $S(x \wedge y) = S(x) \cap S(y)$ and therefore $\mathcal{F} = \{S(x) : x \in L\}$ is intersection-closed and $|\mathcal{F}| = |L|$. Moreover, the size of the universe of $\mathcal{F}$ is at most n, so we may assume that $\mathcal{F}$ has a rare element x, then clearly x is join-irreducible and is contained in at most half of the sets in $\mathcal{F}$. Then for any $y \succcurlyeq x$ clearly $x \in S(y)$ and thus the number of elements greater than or equal to x is bounded by the number of sets in $\mathcal{F}$ containing x, which is at most $\frac{1}{2}|L|$. $\square$

Corollary 2.7.2.1: *Every finite lattice $(L, \preccurlyeq)$ with at least two elements and at most 14 join-irreducible elements has a join-irreducible element e such that at most half of the elements of L are greater than or equal to e.*

Additionally, we restrict the graph formulation 1.4.5, proof of which can be found in [5] and [4], and obtain a result for bipartite graphs.

Observation 2.7.3: *If the Union-closed sets conjecture holds for all finite union-closed families on at most n elements, then the graph formulation 1.4.5 holds for all bipartite graphs with each bipartition class containing at most n vertices.*

Proof: Let G be a bipartite graph with partition classes X, Y and let the size of each X, Y be at most n. By symmetry, it is sufficient to find a rare vertex in X. Denote $\mathcal{F}$ the set of all maximal stable sets in G and define the family $\mathcal{H} := \{A \cap X : A \in \mathcal{F}\}$ by intersecting member sets of $\mathcal{F}$ with the partition class X. It is straightforward to check that $\mathcal{H}$ is intersection-closed. We note that $U(\mathcal{H}) = X$, since X is a stable set in G. Assume that all intersection-closed families on at most n elements have a rare element, then $\mathcal{H}$ must have a rare element e, since $|U(\mathcal{H})| = |X| \leq n$. It is immediate that e belongs to at most half of the maximal stable sets in G and that $e \in X$, which completes the proof. □

Corollary 2.7.3.1: *Any bipartite graph G with at least one edge. If the smaller bipartition class of G contains at most 14 vertices, then G contains in each of it's bipartition classes a vertex that lies in at most half of the maximal stable sets of G.*

Moreover, in Observation 2.7.2 and Observation 2.7.3 we can replace the assumptions of Theorem 2.7.1 with the assumptions of Corollary 2.7.1.1 and get new results. The proofs are otherwise identical.

Corollary 2.7.3.2: *The lattice formulation holds for all lattices of size at most 58.*

Corollary 2.7.3.3: *The graph formulation holds for all bipartite graphs which have at most 58 maximal stable sets.*

3 Conclusion

In this thesis, we studied the Union-closed sets conjecture and related partial results. We discussed the assumptions of the conjecture, it's equivalent formulations, Frankl-complete families, recent breakthroughs and results regarding small families of sets.

We showed that the Union-closed sets conjecture holds for all union-closed families on at most 14 elements, which is a new result improving upon the previously known bound 12 by Vučković a Živković [28] in 2017. Our proof is computer assisted. We formulate a restriction of the Union-closed sets conjecture with a bound on the biggest set as an integer linear programming problem. This problem is then relaxed and to compensate for the relaxation we split the proof into multiple cases and introduce stronger constraints. The whole proof can be readily verified through a provided verification program using linear programming duality.

Our proof of the restriction of the Union-closed sets conjecture has corollaries in terms of lattice and graph theory. We showed that every bipartite graph with each bipartition class containing at most 14 vertices contains a vertex in each of it's bipartition classes which belongs to at most half of the maximal stable sets. We also showed that all lattices L with at most 14 join-irreducible elements contain a join-irreducible element x such that at most half of the elements of L a greater than or equal to x.

We believe that our results can be further improved using the methods from this thesis. Increasing the bound in the restricted conjecture will result in each case taking longer to compute and have lower objective value, which can be balanced by stronger constraints and a faster computer.

References

[1] R. Alweiss, B. Huang, and M. Sellke: *Improved Lower Bound for the Union-Closed Sets Conjecture*, arXiv:2211.11731, (2022).

[2] I. Balla, B. Bollobas, and T. Eccles: *Union-closed families of sets*, J. Combin. Theory (Series A), vol. 120, pp. 531–544, (2013).

[3] I. Bošnjak and P. Marković: *The 11-element case of Frankl's conjecture*, European Journal of Combinatorics, vol. 15, R88, (2008).

[4] H. Bruhn and O. Schaudt: *The journey of the union-closed sets conjecture*, Graphs and Combinatorics, vol. 31, pp. 2043–2074, (2015).

[5] H. Bruhn, P. Charbit, O. Schaudt, and J. Telle: *The graph formulation of the union-closed sets conjecture*, European Journal of Combinatorics, vol. 43, pp. 210–219, (2015).

[6] S. Cambie: *Better bounds for the union-closed sets conjecture using the entropy approach*, arXiv:2212.12500, (2022).

[7] S. Cambie: *Progress on the union-closed conjecture and offsprings in winter 2022-2023*, arXiv:2306.12351, (2023).

[8] Z. Chase and S. Lovett: *Approximate union closed conjecture*, arXiv:2211.11689, (2022).

[9] V. Chvátal: *Linear Programming.* W. H. Freeman, (1983).

[10] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver: *Combinatorial Optimization.* John Wiley & Sons, (1998).

[11] G. Lo Faro: *Union-closed sets conjecture: improved bounds*, Journal of Combinatorial Mathematics and Combinatorial Computing, vol. 16, pp. 97–102, (1994).

[12] G. Lo Faro: *A note on the union-closed sets conjecture*, J. Austral. Math. Soc., Series A, vol. 57, pp. 230–236, (1994).

[13] P. Frankl: *Extremal set systems*, in Handbook of combinatorics, Elsevier, (1995), pp. 1293–1329.

[14] J. Gilmer: *A constant lower bound for the union-closed sets conjecture*, arXiv:2211.09055, (2022).

[15] Gurobi Optimization LLC: *Gurobi Optimizer Reference Manual.* 2023. [Online]. Available: https://www.gurobi.com/

[16] E. Knill: *Graph generated union-closed families of sets*, arXiv:math/9409215v1, (1994).

[17] F. Marić, B. Vučković, and M. Živković: *Fully Automatic, Verified Classification of all Frankl-Complete (FC(6)) Set Families*, arXiv:1902.08765, (2019).

[18] P. Marković: *An attempt at Frankl's conjecture*, Publications de l'Institut Mathématique, vol. 81, pp. 29–43, (2007).

[19] L. Pebody: *Extension of a Method of Gilmer*, arXiv:2211.13139, (2022).

[20] B. Poonen: *Union-closed families*, Journal of Combinatorial Theory, Series A, vol. 59, pp. 253–268, (1992).

[21] R. Morris: *FC-families and improved bounds for Frankl's conjecture*, European Journal of Combinatorics, vol. 27, pp. 268–282, (2006).

[22] J. Reinhold: *Frankl's conjecture is true for lower semimodular lattices*, Graphs and Combinatorics, vol. 16, no. 1, pp. 115–116, (2000).

[23] J. Renaud and D. Sarvate: *On the union-closed sets conjecture*, Ars Combinatorica, vol. 27, pp. 149–154, (1989).

[24] I. Rival: *Graphs and order*. Springer Netherlands, (1985).

[25] W. Sawin: *An improved lower bound for the union-closed set conjecture*, arXiv:2211.11504, (2022).

[26] W. A. Stein and others: *Sage Mathematics Software (Version 10.3)*. 2024. [Online]. Available: http://www.sagemath.org/

[27] T. P. Vaughan: *Families Implying the Frankl Conjecture*, European Journal of Combinatorics, pp. 851–860, (2002).

[28] B. Vučković and M. Živković: *The 12-element case of Frankl's conjecture*, IPSI BgD Transactions on Internet Research, vol. 13, pp. 65–71, (2017).

[29] P. Wójcik: *Density of union-closed families*, Discrete Mathematics, vol. 105, pp. 259–267, (1992).